SUBTASK - 3

Given: For Gaussian discriminant analysis, joint probability distribution is given by:

$$p(y) = \begin{cases} \phi & \text{if } y=1 \\ 1-\phi & \text{if } y=0 \end{cases}$$

$$p(x|y=0) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1}(x-\mu_0)\right)$$

$$p(x|y=1) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_1)^T \Sigma^{-1}(x-\mu_1)\right)$$

where ϕ, μ_0, μ_1 and Σ are parameters of the model.

T.P.T: $$p(y=1|x;\phi,\mu_0,\mu_1,\Sigma) = \frac{1}{1+\exp(-(\theta^T x+\theta_0))}$$

where $\theta \in \mathbb{R}^n$ and $\theta_0 \in \mathbb{R}$.

Proof:

By Bayes' Rule, which states that,
if $B_1, B_2 \dots B_n$ partition the sample space S and if A is an event with $P(A) > 0$, then for $j = 1,2 \dots n$, we have,

$$P(B_j/A) = \frac{P(A/B_j)\,P(B_j)}{P(A)}$$

Applying Bayes' Rule, we get that,

$$p(y=1/x;\phi,\mu_0,\mu_1,\Sigma) = \frac{P(x/y=1)\,P(y=1)}{P(x)}$$

Note: Omitting formal notation for now, eg $P(x/y=1;\phi,\mu_0,\mu_1,\Sigma)$

From law of total probability,
(using the same notation), $P(A) = P(A/B_1)P(B_1) + \dots\dots + P(A/B_n)P(B_n)$

$$\therefore P(x) = P(x/y=1).P(y=1) + P(x/y=0)\,P(y=0)$$

Substituting in earlier eq^n, we get,

$$P(y=1/x;\phi,\mu_0,\mu_1;\Sigma) = \frac{P(x|y=1).P(y=1)}{P(x|y=1)P(y=1)+P(x|y=0)P(y=0)}$$

$$= \frac{1}{1+\frac{P(x|y=0)P(y=0)}{P(x|y=1)P(y=1)}}$$

$$\frac{P(x|y=0)\ P(y=0)}{P(x|y=1)\ P(y=1)} = \text{substituting formulae given} = \left(\frac{1-\phi}{\phi}\right)\frac{\exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)}{\exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)}$$

$$= \left(\frac{1-\phi}{\phi}\right)\exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)+\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)$$

Note: $\underbrace{(x-\mu_0)^T}_{(1\times n)}\ \underbrace{\Sigma^{-1}}_{(n\times n)}\ (x-\mu_0)$ = (1×1) dimensional = scalar

Ↄ let this be $(n\times 1)$ dimensional

As it is a scalar, it can also be expressed as $\frac{(x-\mu_0)^2}{\Sigma}$.

$$\frac{P(x/y=0)\ P(y=0)}{P(x/y=1)\ P(y=1)} = \left(\frac{1-\phi}{\phi}\right).\exp\left(-\frac{1}{2}\frac{(x-\mu_0)^2}{\Sigma}+\frac{1}{2}\frac{(x-\mu_1)^2}{\Sigma}\right)$$

$$= \left(\frac{1-\phi}{\phi}\right)\exp\left(\frac{-1}{2\Sigma}(x^2-2\mu_0 x+\mu_0^2-x^2+2\mu_1 x-\mu_1^2)\right)$$

$$= \left(\frac{1-\phi}{\phi}\right)\exp\left(\frac{-1}{2\Sigma}((2\mu_1-\mu_0\times 2)x+\mu_0^2-\mu_1^2)\right)$$

$$= \exp\left(-\left(\frac{\mu_1-\mu_0}{\Sigma}\right)x-\left(\frac{\mu_0^2-\mu_1^2}{2\Sigma}-\log_e\left(\frac{1-\phi}{\phi}\right)\right)\right)$$

$\therefore$ We observe that, $P(y=1/x;\phi,\mu_0,\mu_1,\Sigma) = \dfrac{1}{1+\exp(-(\theta^T x+\theta_0))}$

where $\theta_0 = \dfrac{\mu_0^2-\mu_1^2}{2\Sigma}-\log_e\left(\dfrac{1-\phi}{\phi}\right)$

$$\theta^T = \frac{\mu_1 - \mu_0}{\Sigma}$$

Hence proved.